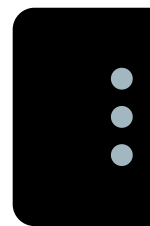


07 Distribuições de frequência com variáveis quantitativas

PRÓXIMA ATIVIDADE



Quando construímos distribuições de frequências com variáveis quantitativas, precisamos, na maior parte dos casos, definir as classes em que desejamos agrupar os nossos dados. O **pandas** nos oferece a função `cut()` para realizar esta tarefa.

Analise o seguinte código:

```
classes = [dados.Altura.min(), 1.65, 1.75, dados.Altura.max()]
labels = ['1 - Baixa', '2 - Média', '3 - Alta']
```

```
frequencia = pd.value_counts(
    pd.cut(
        x = dados.Altura,
        bins = classes,
        labels = labels,
        include_lowest = True
    )
)
```

```
percentual = pd.value_counts(
    pd.cut(
        x = dados.Altura,
        bins = classes,
        labels = labels,
        include_lowest = True
    ), normalize = True
) * 100
```

```
dist_freq_altura = pd.DataFrame(
    {'Frequência': frequencia, 'Porcentagem (%)': percentual}
)
```

```
dist_freq_altura.rename_axis('Estaturas', axis= 'columns', inplace= True)
```

```
dist_freq_altura.sort_index(ascending = True, inplace = True)
```

```
dist_freq_altura
```

[COPIAR CÓDIGO](#)

Utilizando a variável **Altura**, do nosso *dataset* de aula, indique a alternativa que contém a distribuição de frequências gerada pelo código acima.

Estaturas Frequência Porcentagem (%)

1 - Baixa	20529	26.716554
2 - Média	36162	47.061426
3 - Alta	20149	26.222020



Alternativa correta! Esse é o resultado do código acima. Para esta classificação, utilizamos valores bem próximos do 1º e 3º quartis (ainda vamos estudar estas medidas). Observe que temos quase 25% das pessoas com classificação baixa, 25% com classificação alta e 50% com classificação média.

Estaturas Frequência Porcentagem (%)

1 - Baixa	20529	0.267166
2 - Média	36162	0.470614
3 - Alta	20149	0.262220

Alternativa errada! Nessa alternativa, foi esquecido de multiplicar a `Series` `percentual` por 100.

Estaturas Frequência Porcentagem (%)

1 - Baixa	20149	26.222020
2 - Média	36162	47.061426
3 - Alta	20529	26.716554

Alternativa errada! Nessa alternativa, a lista `labels` foi criada de forma errada:

```
labels = ['3 - Alta', '2 - Média', '1 - Baixa']
```

[COPIAR CÓDIGO](#)**Estaturas Frequência Porcentagem (%)**

1 - Baixa	20528	26.715600
2 - Média	36162	47.062039
3 - Alta	20149	26.222361

Alternativa errada! Aqui, o parâmetro `include_lowest = True` foi esquecido.